DOCUMENT RESUME

ED 123 576 CS 002 658

AUTHOR Arlin, Marshall
TITLE Causal Priorities between Comprehension Subskills: Word Meaning and Paragraph Meaning.
PUB DATE 76
NOTE 24p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, April 19-23, 1976)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Definitions; Junior High Schools; *Paragraphs; *Reading Comprehension; *Reading Processes; Reading Research; *Remedial Reading; Remedial Reading Programs

ABSTRACT

Causal priorities between word meaning ability and paragraph meaning ability were examined in three types of readers across a period of nine months. Subjects were 180 pupils in seventh and eighth grade. They were categorized by relative performance in word and paragraphs meaning into three groups defined as "balanced," and "word dominant," and "paragraph dominant." Cross-lagged correlation analysis supported the causal priority of word meaning in balanced and word dominant readers and a different pattern in the paragraph dominant readers. For most readers, the conventional wisdom hypothesis of word meanings as building blocks of paragraph meaning received empirical support. (Author)

Causal Priorities between Comprehension Subskills: Word Meaning and Paragraph Meaning*

Marshall Arlin
University of British Columbia

Running head: Causal Priorities between comprehension subskills

*Paper presented to AERA SIG: Basic Research in Reading. April 1976, San Francisco. Submitted for publication. Author's address: Marshall Arlin, Ph.D. Dept. of Educational Psychology, Faculty of Education. Univ. of B.C. Vancouver, B.C. Canada V6T 1W5

Abstract

Causal priorities between word meaning ability and paragraph meaning ability were examined in three types of readers across a period of nine months. Subjects were 180 pupils in grades 7 and 8. They were categorized by relative performance in word and paragraph meaning into three groups (n=60) defined as "balanced", "word dominant" and "paragraph dominant". Cross-lagged correlation analysis supported the causal priority of word meaning in balanced and word dominant readers and a different pattern in the paragraph dominant readers. For most readers, the conventional wisdom hypothesis of word meanings as building blocks of paragraph meaning received empirical support.

Margaret Early, former president of NCTE has stated recently that "Obviously the next major goal of reading research must be unraveling the mysteries of comprehension.... We still know next to nothing about the process, even though we are continually devising new measures of whatever it is." (Early, 1976). The present study represents a primitive start at unravelling some of these mysteries by empirically examining a tenet of conventional wisdom concerning comprehension. This tenet is the belief that, at the very least, comprehension involves two inter-related but distinct skills, (aspects, processes, etc.)- word meaning and paragraph meaning. Further, of the two skills, word meaning is the more basic. Word meaning appears to be a necessary but not sufficient cause of paragraph meaning. A reader cannot derive meaning from a paragraph unless he can derive meaning from the individual words in the paragraph. Perhaps because this tenet is so intuitively appealing and "obvious", it appears to be taken axiomatically, and few, if any, attempts have been made to verify it. Although comprehension has been analyzed into exceedingly complex sets of interrelationships and hierarchies among subskills, the causal inter-relationship between word meaning and sentence/paragraph meaning appears to have been taken for granted. In the present study, I suggest an approach which may allow examination of this, and other causal inter-relationships among subskills of comprehension.

Many researchers include word meaning as a major component of comprehension (Carroll, 1972; Davis, 1972; Posner, Lewis and Conrad, 1972; Thurstone 1946). There is less agreement about other aspects of comprehension, and in particular about paragraph meaning as a factor distinct from word meaning. Davis (1944, 1972), for example, has posited six or seven comprehension subskills in addition to word meaning but Thurstone (1946) reanalyzed Davis' earlier data and concluded there were only two factors, word meaning and "another factor". Carroll (1972) analyzed comprehension into memory for word meaning and three or four other aspects. Farr (1969) has argued that the distinction between 'vocabulary' and 'comprehension', so common to test constructors and publishers is without much research evidence. With Farr, I am dissatisfied with the distinction, but perhaps on different grounds. "Vocabulary" tests which measure word meaning should not be distinguished from "comprehension" tests measuring paragraph meaning, since I suggest they both measure different aspects of comprehension.

For the present argument, I assume that comprehension requires, in addition to word meaning, the ability to process sentence or paragraph meanings. I further assume that these two separable subskills can be measured with reasonable accuracy by certain standardized reading tests. Having made these assumptions, I ask two questions in this study. First, is there a causal relationship between word meaning and paragraph meaning? Second, does this relationship vary between different types of readers?

It is proposed that in normal circumstances, word meaning is causally prior to paragraph meaning. "Causally prior" is a technical phrase used in social science theory construction (e.g. Kenny, 1973, in press) to indicate that with two correlated variables, one variable is a better predictor of its correlate than the other. Unlike most correlational techniques which do not permit causal inference, this technique does allow tentative causal inference (see Statistical Analysis, below).

In normal circumstances, word meaning is assumed to influence ("cause") performance in sentence or paragraph meaning more than paragraph meaning influences word meaning. If this were true, word meaning could be described as causally prior to paragraph meaning. I have used the phrase, "under normal circumstances" to indicate that the causal priority of word meaning is predicted for most, but not all readers. Rather than hypothesizing a general pattern of causal priority, it seems plausible to hypothesize distinct causal priorities, based upon the reading group under consideration. In this study, I suggest three categories of readers, and hypothesize a pattern of causal priority for each group.

Readers are categorized according to the observed difference between ability in word meaning and ability in paragraph meaning. This categorization technique is similar to techniques used by others who have compared differences between reading subskills (e.g. Guthrie, 1973; Ohaver, 1972; Weiner and Cromer, 1967).

The first type of reader is called balanced. This reader's skills in both word meaning and paragraph meaning are roughly equivalent (hence "balanced"). If he is poor in one skill, he is poor in the other; if he is excellent in one, he is excellent in the other. This is assumed to be the normal case. This type of reader is hypothesized to exhibit a moderate causal priority of word meaning. For this type of reader, word meanings are the building blocks of paragraph meaning. Ability to derive meaning from words is more influential upon ability to derive meaning from paragraphs than vice versa.

The second type of reader is called word dominant. This reader's skill in word meaning is considerably greater (hence "dominant") than his ability in paragraph meaning. If he is superior in word meaning he would be average or worse in paragraph meaning. If he is below average in word meaning, he would be abysmal in paragraph meaning. This type of reader is considered to be a deviation from the normal pattern. The observed difference in abilities is hypothesized to reflect a cumulative deficit in paragraph meaning ability. This deficit reflects an abnormal causal priority of word meaning which has typified this reader's growth in both subskills. His growth is slower in paragraph meaning than word meaning because he is overly dependent upon word meanings as building blocks for paragraph meaning. He cannot 'fill in' paragraph meanings in text involving unfamiliar words as easily as the balanced reader. This reader should be characterized by an exaggeration of word meaning causal priority, reflecting his over-dependence upon word meaning.

The third type of reader is called paragraph dominant. This reader's skill in paragraph meaning is considerably greater than his ability in word meaning. This type of reader also is considered to be a deviation from the normal pattern. The observed difference in abilities is hypothesized to represent a cumulative deficit in word meaning ability. This deficit reflects a causal priority of paragraph meaning which has typified this reader's growth in both subskills. Lack of knowledge of individual word meanings in a paragraph does not hinder his ability to understand the whole paragraph as much as might be expected in the normal situation. For this reader, the whole is greater than the sum of its parts. For this type of reader there is hypothesized to be a causal priority of paragraph meaning. Ability in paragraph meaning should influence growth in word meaning more than ability in word meaning should influence growth in paragraph meaning. Whereas the word dominant reader might be said to understand the whole of a paragraph through its parts (words) the paragraph dominant reader might be said to understand the parts of a paragraph through its whole.

In none of the types of readers is it assumed that one process operates exclusively of the other. With all three types of readers it is assumed that the two subskills are interactive - word meanings facilitate paragraph meanings and vice versa. In all three types of readers, it is assumed that the common factor variance between the two subskills is considerably greater than the unique factor variance attributable to either subskill. However, it is the unique aspects of each (the

differences between the subskills) which are examined here in the hopes that they will permit analyses of the hypothesized causal priorities.

The causal priorities are now proposed as expected patterns to be examined in an empirical manner. Since the study is exploratory and the number of subjects small, these are not to be construed as formal hypotheses to be subjected to statistical tests. The data from this initial study will be examined to see if the observed patterns are similar to those which might be expected if different causal paths typify the three groups. These expectations are:

1) Balanced readers will exhibit a pattern of causal priority of word meaning in the acquisition of new word meaning and paragraph meaning skills.

2) Word dominant readers will also exhibit causal priority of word meaning in the acquisition of new word meaning and paragraph meaning skills in a pattern even more marked than that exhibited by balanced readers.

3) Paragraph dominant readers will exhibit a pattern of causal priority of paragraph meaning in the acquisition of new word meaning and paragraph meaning skills.

Method

Variables and Their Measurement

The Gates-MacGinitie vocabulary and comprehension tests (Form D, Levels 1 and 2, Grades 4-6) were used as pre and post measures of word meaning and paragraph meaning respectively. Use of the terms word meaning and paragraph meaning to represent unitary dimensions is simplistic, and for precision it is necessary to specify each of these terms as operationally restricted by use of the Gates-MacGinitie test. The 'vocabulary' test appears to measure recognition of semantic synonyny (Katz, 1973). The student is presented a stimulus word plus five response alternatives and is told "Find the one word in the group below it that means most nearly the same..." The comprehension test appears to measure recognition of semantic anomaly, (Katz 1973) through a modified cloze procedure. For example, following a sentence about Lindbergh's nonstop flight to Paris comes the sentence, "Jet planes now __25__ the Atlantic Ocean take only a __26__ of the time that Lindbergh took." Five choices are given for each of #25 and #26. This sentence in part assesses paragraph meaning in the sense of deriving meaning from connected discourse, but also appears to require the pupil to recognize that some of the choices while syntactically feasible are semantically anomalous. (e.g. "...jet planes now refueling the

Atlantic Ocean...") Thus in an analytic, but jargon-laden sense, a more appropriate title for the present study might be "The Causal Priorities between Synonymy and Semantic Anomaly". Even this would not be a completely accurate title because some of the 'comprehension' items are dependent upon syntactic cues as well as semantic cues, and some of the 'vocabulary' items are not synonymies but are super-ordinate-subordinate relations. In addition the words used in both exercises are different which allows observed differences between scores on the two tests to be influenced by sampling errors. For purposes of the study, I assume the two subtests validly and reliably sample from the respective domains of word meaning and paragraph meaning. It is necessary to note that the terms word meaning and paragraph meaning are quite specific and operationally restricted by the subtests used in the study.

Statistical Analysis

The analysis was a post hoc cross-lagged panel correlation (Campbell and Stanley, 1963 and Andrews, 1964; Kenny, 1973). According to this technique, panel data can be analyzed to indicate which of two variables A and B, each measured at time 1 and 2 is more likely to have causal priority over the other. If A determines B rather than the reverse, then the cross-lagged correlation A_1B_2 should exceed B_1A_2.

The type of subject and treatment is incidental to the analysis, and to the argument presented. The reason for choosing the particular subject sample and treatment was accessibility to this

data and was not determined by any theoretic aspect of the rationale. Nevertheless, both the subjects and the treatment will be described below. The subjects are described because they may influence the generalizeability of the results. The treatment is described because it may influence the generalizeability and because it satisfied one prerequisite of cross-lagged analysis - it ensured that a significant change took place during the time lag between measurements.

Subjects

Subjects were 180 pupils drawn from a sample of approximately 600 grade 7 and 8 pupils participating in a remedial reading program (described below). Pupils were primarily Black males of lower and lower-middle SES in a Southeastern U.S. city with a population of 200,000. The 600 pupils were selected for the program using the criterion of reading scores below a grade level equivalent of 5.0 on the combined vocabulary and comprehension tests of the Gates-MacGinitie reading survey. The screening test also served as the pre-test for those who participated in the remedial treatment. After completion of the program, and on the basis of the pre-test, pupils were categorized by type of dominance.

Word dominant readers were defined as those whose pre-test grade equivalent vocabulary scores were .9 years higher than their pre-test grade equivalent comprehension scores. Balanced readers were those whose scores were within plus or minus .2 years of each other. Finally, paragraph dominant readers were defined as those whose comprehension scores

were at least .9 years higher than their vocabulary scores. At the end of the treatment, there were complete pre-post data for 479 pupils of whom 68 were comprehension dominant, 98 were balanced and 62 were vocabulary dominant. From each of these groups 60 pupils were randomly selected for analysis.

Treatment

Although the study is non-experimental, and only indirectly related to the treatment, the sample of the study was used in the expectation that as a result of the treatment, subjects would undergo significant increase on at least one of the variables. The treatment was one hour per day of special reading classes during a nine month school year. Each class was conducted by a reading specialist. The reading classes included options such as adult reading tutors, word games like "Password", programmed instruction, high interest-low vocabulary paperback novels and a variety of audio-visual materials. The reading activities stressed both vocabulary building and comprehension development.

The treatment differed both between and within classes, so that treatment for all subjects was not uniform. However the major requirement for the present study was that the treatment (in all its variety) produce significant gains. The usefulness of the cross-lagged correlation technique depends on change that exceeds random fluctuations due to measurement error. Because of the intensive nature of the reading treatment, significant gains were expected on

each of the measures, vocabulary and achievement. These gains did occur and their magnitude and variability were believed to provide adequate opportunity to detect causal priorities.

Results

Pre and Post test results of the mean scores for each of the three groups are presented in Table 1.

Insert Table 1 about here

As anticipated, there was a significant increase in each group on both variables (all correlated t-tests were significant beyond the .05 level). More importantly, the correlation networks appeared to suggest the causal priorities expected for each of the three groups, as shown in Figure 1.

Insert Figure 1 about here

Although the differences in cross-lagged correlations (V_1C_2 vs. C_1V_2) are not great, they are all in the expected directions. Further, the rank order causal priorities of the three sets of correlations corresponds to that expected. The balanced readers exhibited a pattern suggesting causal priority of word meaning. The word dominant readers exhibited this pattern, but to a slightly stronger degree, as anticipated. Finally, while the difference in the cross-lagged correlations of paragraph

dominant readers was miniscule (.64 vs. .60) it was in the expected direction, and is quite different from the pattern suggested by the other two groups.

Alternate Explanation of Results

Considering that the study is exploratory and the number of subjects small, it is desirable to consider some competing hypotheses which might account for the observed patterns. One alternate hypothesis might explain the results in terms of a path through the stability coefficient of the relatively higher comprehension score coefficient. In all three cases, the horizontal correlations (V_1V_2 or C_1C_2) representing stability coefficients suggest a greater degree of reliability or stability for the comprehension measure. However, since this is equally the case in all three groups, this should not cause the differences in patterns which were observed. In all three groups, an argument could be made for a path through the stronger of the two stability coefficients to explain the strength of the V_1C_2 correlation. For example in the balanced group the V_1-.99-C_1-.78-C_2 path might partially account for V_1--.76--C_2. This path is stronger than the C_1--.99--V_1--.64--V_2 path which could be said to account for the C_1--67--V_2 cross-lagged correlation. However, the paths C_1--C_2--V_2 and V_1--V_2--C_2 also capitalize on passing through the stronger of the two horizontal correlations and might just as reasonably be used to explain the cross-lagged correlations. In the paragraph dominant group, the pattern is reversed. Moreover, even if a particular path were confounding the differences in cross-

lagged correlations, it would be difficult to explain the reversal of pattern in paragraph dominant pupils.

Another explanation of the results might involve a possible regression of coefficients toward the population mean in all three groups. I obtained a V_2C_2 coefficient of .66 for 479 pupils (including the 180 in the present sample), and the Gates-MacGinitie manual reports vocabulary-comprehension coefficients in the mid .70's. Therefore the "true" coefficient (i.e. population coefficient) in the .70's seems reasonable. The initial coefficients should be nearly perfect because 'balanced' pupils were by definition almost identical on both measures. In the balanced case, one expects the initial coefficient to be spuriously high, and over time for it to move toward a population mean which is probably in the .70's. This pattern can be observed in Figure 1. The opposite pattern might be expected in the other two groups - pupils chosen because their two scores were disparate. Nevertheless, for both dominant groups a pattern of increasing vertical coefficients across time is not noticeable. The vertical coefficients of paragraph - dominant pupils appear to approximate the population coefficient. The word dominant group is peculiar. The initial V_1C_1 correlation is higher than would be expected for a group selected on the basis of its disparateness and it decreases across time, to a level somewhat lower than one would expect as an estimate of the population coefficient. Although regression is undoubtedly taking place, it

does not seem to explain the observed patterns.

Another regression pattern is possible, suggested by the data presented in Table 1. It will be noted in Table 1 that paragraph dominant and word dominant readers become more 'balanced' over time. One possible explanation for this is an interaction with the treatment such that consciously or unconsciously, pupils work on (or are encouraged by teachers to work on) that aspect of reading in which they are most deficient, either word meaning or paragraph meaning. This might account for the observed regression toward balance.

A third type of regression hypothesis requires that the pre-treatment differences be artifacts of selected extremes. Regression would then be toward the 'true', balanced state of pupils, who were erroneously screened into the two non-balanced categories. This type of regression, if it exists, does not lessen the argument for the respective priorities observed, since in each case the higher of the two scores exhibits causal priority during the return to the state of 'balance'. If the imbalanced conditions reflect 'true' differences in reading style and skill acquistiion, one would expect that the imbalances would remain constant over time. Thus the word dominant pupil should continue to exhibit the same degree of word meaning dominance both before and after the treatment.

Discussion

Results should be interpreted in a cautious and tentative manner, bearing in mind the exploratory nature of the study. In this initial attempt, the results in two cases conformed quite closely to those which were anticipated. The results appear to offer support for the conventional wisdom of considering word meanings to be the building blocks of paragraph meanings. This was found to be the case for balanced readers and word dominant readers. Allowing for differences in operational definitions of difference and for tests with different standard errors of measurement, a rough estimate of balanced readers might include roughly 60% balanced and 20% word dominant. Therefore roughly 80% of the reading population could be hypothesized to follow a pattern of word meaning priority.

What about the other 20%, the paragraph dominant readers? Even though the results for this group were the most open to question, I consider the results from this group the most intriguing.

Even though the pattern for this group did not conform to the paragraph priority expected, the lack of causal priority may in itself suggest a particular mode of acquiring reading skills.

Aware of the fallacy of 'confirming the null hypothesis', I suggest that this particular null pattern is surprising

and partially supports the expectations for these readers. It is intuitively appealing to find results which support the causal priority of vocabulary in the other two types of readers. After all, what could be more sensible than a words-as-building-block hypothesis? However, it was suggested that the causal pattern for paragraph dominant readers would be the opposite of this sensible pattern. Although paragraph dominant readers did not reverse the pattern, they appear to have run counter to the 'sensible' pattern. It is difficult to understand how a group of readers can fail to exhibit causal priority of vocabulary and for this reason, I suggest that comprehension dominant readers may be particularly worthy of further investigation.

The observed patterns appear to suggest that further research along the present lines might be profitable. Replications across normal readers with different time lags and grade levels is suggested. Further replications should increase the sample size (e.g. 1,000+ per group) to allow use of statistical tests such as the Pearson-Filon (cited in Kenny, 1973). Given the unobtrusive, ex post facto, nature of the approach, a person with access to reading scores of for example 10,000 pupils at two different times could replicate the present study. Even increasing the difference criterion to 1.3 or more years difference, he could still secure more than a thousand pupils per group, thereby increasing the possibility of sampling truly dominant pupils and also increasing the power of the analysis.

Further research crossing the three reading types with types of treatment (word meaning emphasis, balanced treatment, and paragraph meaning emphasis) might also reveal interactions which might suggest relative effectiveness of various treatments for various types of pupils (particularly those who are paragraph dominant).

Finally, future research would profit from using the same vocabulary in the word meaning test as the paragraph meaning test. This would allow more control over the domain-sampling problem in securing readers who were truly dominant or balanced. Since this approach would not be suited to a post hoc analysis of existing data, it might be best to concentrate such a study on that small group whose results appear to defy the norm, the paragraph dominant readers.

In summary, it was hypothesized that causal relationships between word meaning and paragraph meaning skills could be examined empirically, and that results would differ by reading group. Empirical analysis did reveal causal priorities, supporting, with an intriguing exception, the causal priority of word meaning.

Moreover, it is hoped that the method employed might suggest one possible approach for researchers investigating the complex interrelationships between various aspects of reading comprehension.

REFERENCES

Campbell, D. and Stanley, J. Experimental and quasi-experimental designs for research on teaching. In N.L. Gage, Handbook of Research on Teaching, Chicago: Rand McNally, 1963.

Carroll, J.B. Defining language comprehension. In J.B. Carroll and R.O. Freedle (Eds.) Language Comprehension and the Acquisition of Knowledge, N.Y.: Halsted Press, 1972, 1-29.

Davis, F.B. Fundamental factors of comprehension in reading. Psychometrika, 1944, 9, 185-197.

Davis, F.B. Psychometric research on comprehension in reading. Reading Research Quarterly, 1972, 7, 628-678.

Early, M. Important research in reading and writing. Phi Delta Kappan, 1976, 57, 298-301.

Farr, R.C. Problems in measuring reading sub-skills. In R. Farr (Ed.) Reading: What Can Be Measured. Newark, Delaware: International Reading Association, 1969, 33-79.

Gates, A.I., and MacGinitie, W.H. Gates-MacGinitie Reading Tests (Survey D, Forms 1 and 2). N.Y.: Teachers College Press, 1965.

Guthrie, J.T. Models of reading and reading disability. Journal of Educational Psychology, 1973, 65, 9-18.

Katz, J.J. Semantic Theory. N.Y.: Harper & Row, 1972.

Kenny, D. Cross-lagged and synchronous common factors in panel data. In A. Goldberger and O. Duncal Structural Equation Models in the Social Sciences. N.Y. Seminar Press, 1973, 153-165.

Kenny, D. Cross-lagged panel correlation: a test for spuriousness. Psychological Bulletin (in press).

Ohaver, A.R. A comparison study of semantic and syntactic cueing by low reading performance college freshmen. In F. Greene (Ed.), Investigations relating to mature reading. Milwaukee: National Reading Conference, 1972.

Pelz, D.C. and Andrews, F.M. Detecting causal priorities in panel study data. American Sociological Review, 1964, 29, 336-348.

Posner, N., Lewis, J. and Conrad, C. Component processes in reading: a performance analysis. In J. Kavanaugh and I. Mattingly (Eds.) Language by Ear and by Eye. Cambridge: MIT Press, 1972, 159-192.

Thurstone, L.L. Note on a reanalysis of Davis' reading tests. Psychometrika, 1946, 11, 185-188.

Wiener, M. and Cromer, W. Reading and reading difficulty: A conceptual analysis. Harvard Educational Review, 1967, 37, 620-643.

Table 1

Means and Differences between Means

For Vocabulary (V) and Comprehension (C)

Scores for Three Types of Readers at Two Times (1 + 2)

	V_1	V_2	C_1	C_2	$Diff_{V_1-C_1}$	$Diff_{V_2-C_2}$
Balanced	3.61	4.29	3.60	4.55	0.01	0.26
Word Dominant	4.67	5.22	3.26	4.83	1.41	0.39
Paragraph Dominant	3.25	5.18	4.73	5.73	-1.48	-0.55

Figure Caption

Figure 1. Cross-lagged product moment correlations for Vocabulary (V) and Comprehension (C) scores with three types of readers at two times (1 and 2).